

# Challenges in Using Mix Methods in Evaluation

Ph.D. Stefan COJOCARU<sup>3</sup>

## Abstract

This article explores the debates between quantitative and qualitative methods in the evaluation process, analyzes the challenges about methodological mix in terms of credibility and validity of data and tools, and the evaluation findings. Beyond the epistemological contradictions, it seems that, in terms of usefulness, the mixing of methods is a practical solution, along with hybrids theories, able to provide information to improve the sufficiency of the program. Mixing methods is also a way to reduce conflict between positivism and constructivism and an opportunity for increasing flexibility that the evaluator has in choosing the most appropriate methods for obtaining information in the assessment process.

## Keywords:

mix methods; credibility; internal validity; external validity; program evaluation.

---

<sup>3</sup> Stefan COJOCARU - Associate Professor, „Alexandru Ioan Cuza” University, Department of Sociology and Social Work, Iasi, Bd. Carol I no. 11, Tel. 0040.744788779, email: stefan.cojocaru@holtis.ro; director of Holt Romania – Iasi, Childhood and Parenting Research Center, Iasi, str. Bistrita nr. 7, bl. B13, sc. B, ap. 3.

## **Introduction**

At first, the evaluation was mainly based on quantitative methods, that were considered objective, stemming from the positivist approach promoted in the social sciences; different reactions occurred on the necessity and usefulness of qualitative research that led to the development and increase of the dispute between quantitative-qualitative. House believes that "we have entered an ecumenical period, and the qualitative techniques are seen as legitimate and the methodological mix as recommended" (House, 1994, p. 241); in this way exceeds the "schism", as House calls it, from the evaluation domain by combining the two types of techniques, quantitative and qualitative. The discussions and debates, sometimes radical, on the mixing of methods are generated by the fact that research methods are dependent by ontological and epistemological assumptions on how the reality is conceived, and by the characteristics to explain and get to know this reality. Therefore, there are various ways to approach these challenges, either from abstract to concrete (from the ontological level to the methodologic one) or vice versa (from particular data to theories). The first of them follow from the beginning the reconciliation of different theoretical assumptions, and choice of methods and techniques for collecting and interpreting data, so as to satisfy the requirements of internal consistency, while the latter approach is aimed in gathering in many different ways data and interpreting it, in order to broaden the universe of knowledge and to generate new interpretations, based on the multiplicity of data and the perspectives approached (Cojocaru, 2007a; Cojocaru, 2007b). Therefore, to achieve a high degree of internal and external validity, some authors recommend achieving a balance and sustaining a reciprocal relationship between theoretical models and methodologies, between paradigms and practice (Greene and Caracelli, 1997, p. 12). Rossi recommends that the dispute between quantitative and qualitative methods to be left to those who finance evaluations, they, in turn, being informed by the evaluators regarding the evaluation strategy and the type of information provided after the evaluation (1994, p. 35).

### **The dispute between Qualitative and Quantitative**

The dispute between qualitative and quantitative methodologies is sustained and fueled by different theoretical paradigms, depending by the ontological and epistemological assumptions on which they are based; there are also an alternative discourse, which raises the possibility of mixing quantitative and qualitative methods, bringing arguments about ways of living and manifestation of reality, the ways in which they can be known, understood and explained, or by the research methods used to collect data and information needed to verify theories or building new ones, in a deductive approach (from

general to particular), or in an inductive one (starting from empirical data and reaching to expression of some regularities). The discourse between rivalries, which states in particular the advantages of a paradigm over another, is fueled by disputes between objective and subjective, generality and particularity, causality and understanding, extern and intern, representativeness and diversity, facts and values, explanation and understanding, one single reality and multiple realities, categories established and emerging etc. In the history of this rivalry, the representatives of quantitative methods were those who attacked and criticized the qualitative methods, and so the qualitative ones had a defensive position. After the appearance of several reports of qualitative evaluation, the evaluators tried to bring as many arguments in the favour of the reliability and validity of these assessment strategies, and to criticize the lack of detailed information and of applicability of the quantitative data. A new wave of arguments appeared for the reconciliation of the rival positions, with initiatives from both sides, by supporting and arguing for a third paradigm which relies on a combination of the methods and theories. The mixing of methods and the hybridization of theories represents strategies needed for overcoming these rivalries and for combining various methods and techniques, which are aimed to capture from different perspectives the same investigated reality.

### **Qualitative methods and credibility**

When using qualitative methods in the evaluation we admit the importance of subjective information and their relevance for in-depth understanding of the program. Datta (1994) believes that to reach an understanding of the program, the qualitative evaluators must incorporate all the available information in a satisfactory interpretation model, which is able to establish the program influences (p.59). One of the challenges of the qualitative evaluations is represented by the insurance of data credibility; credibility is one of the targets of critics brought by the representatives of the quantitative methods to qualitative approaches in the evaluation process, arguing that the subjective factors alters the interpretation, in terms of proving the influence of the program. If in the quantitative evaluations we speak about the problem of attribution, and this is approached by using quasi-experimental and experimental designs, in qualitative evaluations is pointed out how the program affects the status of the beneficiaries and the environment in which it runs. The evaluation which is based on qualitative methods aims to identify relevant variables, to collect in-depth information from participants, taking into account the context in which this process occurs, and aims to identify the interests, intentions, motives, values and the participants explanations, and it aims to explain the meanings of social realities that are constructed by the participants. The qualitative evaluation exceeds the level of institutional and structural analysis,

based on objective data by querying the meanings of the individual actions in specific or private contexts, therefore, these methods provide detailed information about possible ways of specific intervention to different types of beneficiaries, and refines the analysis of the contexts in which they could produce intentionally changes in the program. For example, quantitative studies conducted in Romania on the issue of "children left home alone", because their parents went to work abroad, want to estimate the number of children affected by this problem, to identify specific features of structural effects and to describe possible solutions to the macro-level; qualitative research aims to identify particular characteristics of how to restructure the family environment, the adaptation strategies of the family in terms of migration and on the different contexts that generate different family realities; from our point of view, the second type of evaluation is more appropriate for identifying the most appropriate and flexible ways of intervention to the specificity of various situations and refining the interpretations concerning the effects upon children, depending on age, gender, existence of phratries, extended family involvement, access to community resources, family background characteristics before departure to work abroad etc. For the application of qualitative methods in the evaluation is necessary to clearly define the limits of the evaluation and to assume objectives that are congruent with ontology and epistemology, which is characteristic to relativism; for example, the qualitative research can not have as a goal to identify causal chains and to argue the interferences that are supporting the theory attribution program effects, these being positivist and post-positivist features of the evaluations: "the theories that depend only by the qualitative methods have little credibility in academic and political circles, whether the results of qualitative research were used to support inferences about the effects achieved for clients or for society in general, following the development of the program (Cook, 2000, pp. 27-28). In addition, using qualitative methods lead to discovery, open new directions for intervention, and refine conclusions from the perspectives of the program participants, and thus, the evaluation.

To ensure the reliability of the data, Raines proposes a series of measures (2008, p. 454): (1) *triangulation of data sources* through the query of at least three categories of actors who can provide information about the program (beneficiaries, involving staff, stakeholders etc. ) (2) *consulting with colleagues* for verifying the construction methods of the sample, methods of collecting data, coding and data analysis procedures, (3) *analyzing the negative cases* necessary to establish the evaluation limits, the topics addressed and the conclusions; (4) *referential adequacy* by ensuring the coherence and exploitation of all sources in a balanced manner, (5) *verifying by the participants* the accuracy of data and the consistency of the conclusions, (6) *keeping the accuracy of the records and transcripts* for ensuring the context of transmission, and the proper interpretation

conditions for those who consult the evaluation report (7) *triangulation of the evaluators* for ensuring the dependability of data, for the purposes of extracting and building the most appropriate codes (8) *testing the hypothesis* and analysing the possible alternative explanations able to provide relevant information about the problem which is studied, or to provide a more appropriate interpretation of it (Saxe and Fine, 1981; Kidder and Judd, 1986; Kidder and Fine, 1997), (9) *theory triangulation* to strengthen a framework for the interpretation of data at different levels of analysis (10) *extending the commitment of the evaluator* to the program through its participation in various activities and ensuring a sufficient period of time to achieve interactions with the participants, thus diminishes the risks relating to the evaluator to be seen as an outsider (11) *triangulation of the methods* can provide different tools for collecting data, with the same epistemological orientation.

### **From the rivalry of discourse to the integration of methods**

Discussions regarding using qualitative methods versus quantitative methods had led to pros and cons arguments, putting in contradiction different science guidelines: objectivism versus subjectivism, realism versus relativism, positivism versus constructivism etc. These conflict debates between these ontological, epistemological and methodological paradigms were analyzed also by Greene and Caracelli (1997), who propose to define and describe the paradigms dispute, which outline and argue the option of mixing the methods in social research, as well as in the evaluation programs. Some qualitative evaluators consider that the quantitative studies produce irrelevant information, and the quantitative accuse the qualitative evaluations as descriptive and not reliable. Thus, each position has highlighted the weaknesses of the rival orientation. Beyond the advantages of mixing methods in evaluation, Chen (1997) believes that quantitative methods can not replace qualitative methods or vice versa, but considers that mixing can not be a dominant strategy in the evaluation, as stated by Datta (1994), because the process of mixing methods can not be considered superior than other methods (for example, by combining different methods is difficult to achieve a higher level of internal validity than if obtained by quantitative methods), and can not build a logical coherence of different theoretical models; on the other hand, Chen (1997) considers that affirming mixing methods, as a dominant strategy of research and evaluation, leads to amplification of rivalries between quantitative and qualitative representatives. On the other hand, appreciating the strengths of each methodological guideline may lead to pragmatic and dialectical combination of methods, subsumed under the goal to bring more knowledge, and so the rivalries can be overcome by combining the perspectives in forming the new generations of evaluators. In the same time, mixing methods represent for Datta

is a step to overcome the rivalries between paradigms, a strategy for solving problems, but also a way of getting others (1994, p. 67). The author considers it as an adaptive paradigm (Datta, 1994, p. 68; Datta, 1997), similar to what House (1994) calls a realistic evaluation. House believes that the quantitative and qualitative methods can be integrated, from the importance of the subject and the need of integrations of the content's outcomes, and the disputes between different paradigms can be overcome through a constructive approach, because of data diversity used in evaluation, such as: facts, figures, descriptions or metaphors (1994, p. 14). Hedrick considers that we do not feel comfortable when we integrate methods, but the way we do it (1994, p. 50). On the other hand, Smith believes that mixing methods must abandon the endless discussions and rivalries between different theoretical models, and selecting different evaluation methods should be directed by the mental models or by the metaphors from the evaluators mind that condition their actions (1997, p. 73). Based on examples of evaluations, Caracelli and Greene believe that mixing methods is used to integrate different paradigms to produce an understanding of various aspects of the phenomena; the authors carried out a classification of integration strategies (1997, p. 23-25): (1) *recurrence strategy*, characterized by a dynamic and continuous interaction between different methods and different paradigms, that influence the use of such methods; in this strategy are used various methods to obtain depth information, and each method aims to refine the way of data collection and information analysis; (2) *incorporation strategy*, which refers to the design that make connections between various methods, that are in a relationship of opposition, but are applied at different times in different stages of evaluation; this strategy may begin using quantitative methods, after which, according to the results obtained, proceed with the application of qualitative methods (Mark Feller and Button, 1997); also some authors recommend to initiate the evaluation by using qualitative methods and, after identifying the categories and indicators, continue with quantitative methods (Wholey, 1987; Bickman and Person, 1990); (3) *holistic strategy*, that uses conceptual frameworks that guide the collection and interpretation of data and appeals to different methods for the complex understanding of the phenomena, and so the tension between the methods is reduced by building the framework and integrative conceptual structures, (4) *transformative strategy*, that focuses on recognizing the importance of values and actions in the evaluation, to highlight the diverse interests and the pluralism of the participants values.

### **Evaluation criteria of the scientific rigor in evaluation**

Each of the two paradigms in the evaluation (quantitative and qualitative) has built its own tests of scientific rigor according to the ontological assumptions and epistemological presuppositions. Thus, for the evaluation that

utilizes quantitative methods for data collection and statistical analysis are used as criteria for scientific validation the following: internal validity, external validity, reliability instruments and the objectivity (Holosko, 2008). For qualitative evaluations have been proposed and consolidated the following equivalent criteria (Guba, 1981, Lincoln and Guba, 2000; Raines, 2008): *credibility, transferability, dependability and confirmability*; these criteria were supported by the qualitative representatives as an response to the attacks of the quantitative representatives, and this system of criteria was sustained as an alternative to that of quantitative studies (Cojocaru, 2010).

<b>Type of paradigm</b>	<b>Positivism and post-positivism paradigms</b>	<b>Costructionism and interpretivism paradigms</b>
<b>Evaluation criteria of the scientific rigor in evaluation</b>	Internal validity	Credibility
	External validity	Transferability
	Fidelity	Dependability
	Objectivity	Confirmability

Table 1 . Evaluation criteria of the scientific rigor in evaluation

### **Mixing methods - solution to reduce conflicts between positivism and constructivism**

The social reality is extremely complex, therefore, its own investigation requires the use of combined methods for capturing as accurate of this complexity. Regardless of the chosen research method, using only one of these, it limits the data collection and their diversity; any method individually used introduce a series of errors, and the perspective of mixing methods aims to reduce them (Greene and Caracelli, 1997) and the clarification of evaluation limits (House, 1994). This argument was used by various authors to support and strengthen the methodological triangulation strategy (Denzin, 1978; Smith and Klein, 1986; Mathison, 1988; Caracelli Greene, 1997), both the positivist and post-positivist, and also followers of constructivism and interpretativism paradigms. Although the use of multiple and diverse methods is regarded as a good idea for evaluation, Greene and Caracelli (1997) considers that it is not necessarily a condition for ensuring scientific criteria (p. 5). Developing

knowledge, broadening the coverage level and of validity of the explanations and deepening various aspects of the program are all topics of reflection for the evaluator; "the primary reason for the research that uses mixing methods is for a better understanding, generating deeper and larger meanings, developing the important knowledge that is claiming respect against a larger sphere of interests and perspectives" (Greene and Caracelli, 1997, p. 7). The debates on mixing methods in evaluation were directed by two major paradigms: the paradigm of interpretivism or constructivism and the post-positivism paradigm or post empiricism (Greene and Caracelli, 1997, p. 5). Combining different research methods overcome the classical logic of rivalry between quantitative and qualitative methods, and Datta (1994) considers that mixing methods will be the future dominant methodology, because combines the strengths of each method (quantitative and qualitative); in the same tone, Chen (1997) affirms that when we combine different methods, their weaknesses are reduced, and this evaluation strategy can help to extend the purpose of evaluation using triangulation of methods.

### **Using contingency in mixing methods**

For mixing the methods and to reduce rivalries between qualitative and quantitative methods Chen proposes a contingency approach to selecting research methods, considering that no method of assessment does meet all needs, and the solution is to choose the most appropriate method for a given context. To support this idea, Chen proposes the conceptualization of the contextual circumstances in three dimensions (1997, p. 63): (1) the evaluation can require the producing of intensive and contextual informations versus extensive and and accurate informations, (2) the evaluator may have increased access to data available or may have limited access to reliable data, (3) the evaluator may have or not accuracy, concerning data that refers to the influence of program on the environment in which runs. From the combination of these dimensions of evaluation context, Chen builds three configurations of the way in which the evaluator can select different research methods and may combine them to ensure the scientific conditions (access to reliable data, strengthening internal and external validity). Mixing methods is regarded as the most appropriate strategy when the contexts of the program's evaluation require accurate and extensive data, the evaluator having access to reliable data that are available and the program is conducted in a closed system (Chen, 1997, pp 64-65). Sometimes, the evaluations do not presents values directly and neither the theoretical models, that guides the collection and analysis processes, the theoretical perspectives are discrete while the evaluations are made using combined methods; and these methods are chosen depending on the context and the evaluation purposes, and the degree of access to relevant data (Caracelli



and Greene, 1997). To reduce rivalries and conflicts between quantitative and qualitative methods, Chen recommends the use of evaluation based on theory, the following considerations (1997, p. 66): (1) evaluation led by theory produces an controlling goal to quantitative and qualitative methods, focusing on the program framework and the program theory at the expense of the methods used; (2) the evaluation based on theory provides a comprehensive framework for mixing methods, overcoming the conflict and the competition between different methods; (3) this evaluation strategy provides the bases of construction and implementation of plans integrated by mixing methods, ensuring consistency of using methods (4) also, the evaluation based on theory represents a justification for mixing methods in evaluation, in ensuring internal and external validity of data and evaluation results. Causality approach in quantitative methodologies can be enriched by a detailed description of phenomena, without producing contradictions between the two approaches.

### **Advantages of methodological mix**

An advantage of using mixing methods is the flexibility that the evaluator has in choosing the most appropriate methods for obtaining information during the evaluation process, this is more useful when the evaluator is faced with multiple requests from the sponsor, the implementer or decision makers. On the other hand, mixing methods provides both statistical data, which are easier to remember and to use into the rhetoric of the decision makers, and qualitative analysis that capture the depth of the studied phenomena, which constitutes a base for the discussions of the experts and practitioners, in refining the adapted interventions to the specific or typical cases. That is why, the methodological mix offers both the opportunity to meet the requirements of the generalization of results (based on the use of probabilistic samples) and strengthening the credibility of using qualitative methods of data collection and their customization in context. One of the requirements of evaluation is related to data credibility and the conclusions generality: "the evaluator is forced to offer both reliability and generality" (Chen, 1997, p. 69); that is why, mixing methods can be an evaluation strategy that helps in meeting those demands of scientific rigor. Mark and Shotland propose, along with triangulation - which justifies the use of different methods to answer the same questions, also other ways of combining methods as follows (1987, p. 98-99): (1) *the complementary goals model* refers to the sequential application of various research methods that has different purposes and functions; for example, it can be used a quasi-experimental design with comparison group to identify the effects of a parenting program and the path analysis to strengthen the evaluation process; (2) *the interpretability strengthening* involves the use of

quantitative and qualitative methods in clarifying and refining the interpretation of the evaluation results; for example, a quantitative method can be used as a primary method, then apply a qualitative method for refining the description, clarification, interpretation; can also be a reverse process, first is applied a qualitative method (the comprehensive interview for identifying the interpretation models of the behaviors and results of intervention), after which you can build a quantitative tool that can provide a analysis that can be more easily communicated and understood; (3) *evaluating the plausibility of risks* to which the validity is subject, which involves the use of a main method of data collection, followed by a method, so-called secondary, which verifies the validity of the first method; for example, the questionnaire can be used in a quasi-experimental design of time series, in which are evaluated the effects of the intervention (the parents counselors) of a program on his results (reducing the risks of child abuse), followed by comprehensive interviews with parents (as a secondary technique necessary to verify the validity of the obtained results by using the questionnaire).

### **Challenges for mixing methods**

Most difficulties in mixing methods are generated by the need to ensure scientific rigor in evaluation (Chen, 1997; Cojocaru, 2009). Chen believes that the flexibility of using methodological mix (seen as an advantage) is accompanied by the difficulties of methodological rigor fulfillment, because the evaluator is forced to continuously monitor these requirements necessary for each method and combining them (1997, p. 70 ); these difficulties are harder to overcome if the evaluation is conducted by methods [method-driven evaluation] because the pragmatic approach of the methodological mix ,increases the possibility of the appearance of some epistemological contradictions, and also increase the rivalries between quantitative and qualitative methods. To reduce the risks related to the methodological mix effects concerning the evaluation, Datta (1997, p. 34-35) proposes a pragmatic approach in combining the different research methods on the following criteria: (1) practical character, (2) contextuality and receptivity, and (3) anticipated consequences of the evaluation. In the same rhetoric there are also the arguments presented by Smith (1994), who considers that theoretical debates about post-positivism and constructivism paradigms, that feeded the rivalries, have neglected practice and how the methods have been applied, therefore, mixing methods is necessary from pragmatic aspects of their use for data collection, from the program's realities. Smith believes that a purist approach of the qualitative representatives, of supporting and strengthening of the incompatibility between qualitative and quantitative methods, has to be overcome by focusing on the strengths of each

one. Caracelli and Greene (1997, p. 22-23) believes that there are three ways of mixing methods: triangulation, complementarity and expansion.

### **Conclusions**

The using of methodological mix for the program evaluation offers several advantages, capable to overcome difficulties and rhetorics in terms of quantitative and qualitative dispute: a) the flexibility that the assessor has in choosing the most appropriate method for obtaining information in the evaluation process; b) provides both statistical data and qualitative analysis; c) provides conditions of generality and credibility of the evaluation conclusions, d) refines the interpretations built on statistical data and qualitative analysis models.

Beyond these advantages, the methodological mix requires that the evaluators have a series of skills and abilities from the two areas of analysis and interpretation, knowing and assuming different epistemologies and also assuming proper research methods, associated with a design that will allow, in a reasonable context, the mixing methods and the hybridization of theories.

## References

- Bickman, L., Person, K.. (1990). Using program theory to describe and measure program quality, in *New directions for evaluation*, vol. 47, pp. 61-72.
- Caracelli, V., Greene, J. (1997). Crafting mixed-method evaluation design, in *New directions for evaluation*, vol. 74, pp. 19-32.
- Chen, H. (1997). Applying mixed methods under the framework of theory-driven evaluations, in *New directions for evaluation*, vol. 74, pp. 61-72.
- Cojocaru, S. (2007a). Strategii de construire a esantionelor calitative utilizate in evaluarea programelor, in *Revista de cercetare si interventie sociali*, vol. 16, pp. 138-151.
- Cojocaru, S. (2007b). Strategii de esantionare cantitativa in procesul de evaluare a programelor, in *Revista de cercetare si interventie sociala*, vol. 19, pp. 115-127.
- Cojocaru, S. (2009). Clarifying the theory-based evaluation, in *Revista de cercetare si interventie sociala*, vol. 26, pp. 76-86.
- Cojocaru, S. (2010). *Evaluarea programelor de asistenta sociala*, Polirom, Iasi.
- Cook, T. (2000). The false choice between theory-based evaluation and experimentation, in *New directions for evaluation*, vol. 87, pp. 27-34.
- Datta, L. (1997). A pragmatic basis for mixed-method design, in *New directions for evaluation*, vol. 74, pp. 33-46.
- Datta, L.E. (1994). Paradigm wars: a basis for peaceful coexistence and beyond, in *New directions for program evaluation*, vol. 61, pp. 53-70.
- Denzin, N.K. (1978). *The research act: a theoretical introduction to sociological methods*, second edition, McGraw – Hill, NewYork.
- Greene, J., Caracelli, V. (1997). Defining and describing the paradigm issue in mixed-method evaluation, in *New directions for evaluation*, vol. 74, pp. 5-17.
- Guba, E.G. (1981). Criteria for assesing the trustworthiness of naturalistic inquiry, in *Educational communication and technology journal*, vol. 29, pp. 75-91.
- Hedrick, T. (1994). The quatitative-qualitative debate: possibilities for integration, in *New directions for program evaluation*, vol. 61, pp. 45-52.
- Holosko, M. (2008). Evaluating quantitative reseach studies, in Grinnell, R., Jr., Unrau, Y., *Social work research and evaluation. Foundations of evidence-based practice*, Oxford University Press, New York, pp. 423-444.
- House, R. (1994). Integrating the quantitative and qualitative, in *New directions for program evaluation*, vol. 61, pp. 13-22.
- Kidder, L, Judd, C. (1986). *Reasearch methods in social relations*, Rinenhart and Winston, New York.
- Kidder, L., Fine, M. (1987). Qualitative and quantitative methods: when stories converge, in *New directions for program evaluation*, vol. 35, pp. 57-75.

- Lincoln, Y., Guba, E. (2000). Paradigmatic controverises, contradictions, and emerging confluences, in Denzin, N., Lincoln, Y., *Handobok of qualitative research*, ed. 2, Sage Poublication, Thousand Oaks, California, pp. 163-188.
- Mark, M., Shotland, L. (1987). Alternative models for the use of multiple methods, in *New directions for program evaluation*, vol. 35, pp. 95-100.
- Mark. M., Feller, I., Button, S. (1997). Integrating qualitative methods in a predominantly quantitative evaluation: a case study and some reflections, in *New directions for evaluation*, vol. 74, pp. 47-59.
- Mathison, S. (1988). Why triangulate?, in *Educational researcher*, vol 17 (2), pp. 13-17.
- Raines, J. (2008). Evaluating qualitative research studies, in Grinnell, R., Jr., Unrau, Y., *Social work research and evaluation. Foundations of evidence-based practice*, Oxford University Press, New York, pp. 445-461.
- Rossi, P.H. (1994). The war between the quals and the quants: is a lasting peace possible, *New directions for program evaluation*, vol. 61, pp. 23-36.
- Saxe, L., Fine, M. (1981). *Social experiments: methods for design and evaluation*, Sage Publications, Newbury Park, California.
- Smith, L., Kleine, P. (1986). Qualitative research and evaluation: triangulation and multimethods reconsidered, in *New directions for evaluation*, vol. 30, pp. 55-71.
- Smith, M. (1997). Mixing and matching: methods and models, in *New directions for evaluation*, vol. 74, pp. 73-94.
- Smith, M.L. (1994). Qualitative plus / versus quantitative: the last word, in *New directions for program evaluation*, vol. 61, pp. 37-44.
- Wholey, J. (1987). Evaluability assessment: developing program theory, in *New directions for evaluation*, vol. 33, pp. 77-92.